A method for recursively generating sequential rational approximations to $\sqrt[n]{k}$

Joe Nance
nance2uiuc@gmail.com
Department of Mathematics
University of Illinois at Urbana-Champaign

December 28, 2013

Abstract

The goal of this paper is to derive a simple recursion that generates a sequence of fractions approximating $\sqrt[n]{k}$ with increasing accuracy. The recursion is defined in terms of a series of first-order non-linear difference equations and then analyzed as a discrete dynamical system. Convergence behavior is then discussed in the language of initial trajectories and eigenvectors, effectively proving convergence without notions from standard analysis of infinitesimals.

1 Introduction and motivation

Consider for a moment the simple recursion

$$\frac{x_{t+1}}{y_{t+1}} = \frac{x_t + 2y_t}{x_t + y_t}. \tag{1}$$

Choose initial values to set it marching on its way towards $\sqrt{2}$. If we choose initial values $x_0 = y_0 = 1$, the recursion (1) gives a sequence of fractions approximating $\sqrt{2}$ whose behavior is summarized in the table below.

t	x_t/y_t	$\approx$
0	1/1	1
1	3/2	1.5
2	7/5	1.4
3	17/12	1.4167...
4	41/29	1.41379...
5	99/70	1.41429...

For reassurance that the recursion generates a sequence that does infact converge to $\sqrt{2}$, employ the following analysis: For some sequence a_t, if $a_t \to L$, then $a_{t+1} \to L$ as well. So suppose that the recursion (1) has limit L. Multiply the top and bottom of the right hand side of (1) by $1/y_t$ and we get

$$\frac{x_{t+1}}{y_{t+1}} = \frac{x_t + 2y_t}{x_t + y_t} \frac{\frac{1}{y_t}}{\frac{1}{y_t}}.$$

Using the previous fact about limits, we have

$$L = \frac{L+2}{L+1}.$$

This gives $L = \pm\sqrt{2}$. In this analysis, we have picked up the unsettling possibility of this recursion generating a sequence of fractions converging to $-\sqrt{2}$. Discussion of which initial values generate such a sequence is withheld momentarily.

2 $\sqrt{k}$ Recursion

It is convenient to consider recursion (1) as an action on a system of first order linear difference equations given by

$$\begin{cases} x_{t+1} = x_t + 2y_t, \\ y_{t+1} = x_t + y_t, \end{cases} \tag{2}$$

which is a discrete dynamical system, or "DDS". Clearly, if we replace 2 by a positive integer $k \in \mathbb{Z}^+$ we obtain a recursion similar in structure to recursion (1) converging to $\pm\sqrt{k}$ for any initial values x_0, y_0. This recursion is

$$\frac{x_{t+1}}{y_{t+1}} = \frac{x_t + ky_t}{x_t + y_t}. \tag{3}$$

This gives a corresponding DDS

$$\begin{cases} x_{t+1} = x_t + ky_t, \\ y_{t+1} = x_t + y_t, \end{cases} \tag{4}$$

which can be represented in matrix form as

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & k \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix}. \tag{5}$$

Any term in a sequence generated by (5) is generalized as

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} 1 & k \\ 1 & 1 \end{pmatrix}^t \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}. \tag{6}$$

Recursion (3) is recovered by taking ratios of terms with equal indices.

3 $\sqrt[n]{k}$ Recursion

A natural question at this point would be "Does there exist a structurally simple recursion similar to (3) generating a sequence of fractions approximating $\sqrt[n]{k}$?", to which the answer is "kind of". In the spirit of the previous analysis, start out with

$$L = \sqrt[n]{k}$$

so that

$$L^n = k.$$

Add L to both sides to get

$$L^n + L = L + k,$$

then factor out an L:

$$L(L^{n-1} + 1) = L + k.$$

Now divide both sides by $L^{n-1} + 1$ to get

$$L = \frac{L + k}{L^{n-1} + 1}.$$

Use the fact about sequences to obtain

$$\frac{x_{t+1}}{y_{t+1}} = \frac{x_t + ky_t}{\frac{x_t^{n-1}}{y_t^{n-2}} + y_t}, \tag{7}$$

which has the corresponding DDS

$$\begin{cases} x_{t+1} &= x_t + ky_t \\ y_{t+1} &= \frac{x_t^{n-1}}{y_t^{n-2}} + y_t. \end{cases} \tag{8}$$

This latter system is not representable as a simple 2×2 matrix with real entries for arbitrary n. This is because there are now *three* different terms in this system: x_t, y_t, and x_t^{n-1}/y_t^{n-2} which is non-linear.

4 Convergence behavior of the DDS given by Π_n

We are now in a position to analyze (8) using matrix methods. Suppose there *did exist* an $n \times n$ matrix, call it Π_n, such that (8) could be represented as

$$\begin{pmatrix} \vec{x}_{t,1} \\ \vec{x}_{t,2} \\ \vdots \\ \vec{x}_{t,n} \end{pmatrix} = \Pi_n^t \begin{pmatrix} \vec{x}_{0,1} \\ \vec{x}_{0,2} \\ \vdots \\ \vec{x}_{0,n} \end{pmatrix}, \tag{9}$$

where $\vec{x}_{a,b}$ denotes the bth entry of the vector corresponding to a. Or equivalently, $\vec{R}_t = \Pi_n^t \vec{R}_0$. But what would the matrix Π_n look like? Consider the properties of (8) which Π_n must capture in order to faithfully represent convergence behavior of (7). First, recognize that the action on (8) needed to arrive at (7) is taking ratios of terms with equal indices, which is equivalent to taking the ratio of successive entries of a vector $\vec{R}_t$. The entries in an arbitrary vector $\vec{R}_t$ must tend toward those of the dominant eigenvector of Π_n, $\vec{\lambda}_d$. So the ratio of any two successive entries in $\vec{\lambda}_d$ should be equal to $\sqrt[n]{k}$. That is, $\vec{\lambda}_{d,i}/\vec{\lambda}_{d,i+1} = \sqrt[n]{k}$ where $1 \leq i \leq n-1$ and $\vec{\lambda}_{d,i}$ is the ith entry of the dominant eigenvector of Π_n. Such an eigenvector looks like

$$\vec{\lambda}_d = \begin{pmatrix} \sqrt[n]{k}^{n-1} \\ \sqrt[n]{k}^{n-2} \\ \vdots \\ 1 \end{pmatrix}.$$

Recovery of the dominant eigenvalue from the dominant eigenvector is had by solving for λ_d in $(\Pi_n - \lambda_d I_n)\vec{\lambda}_d = \vec{0}$ where $\vec{0}$ is the column vector consisting entirely of zeros. This calculation yields $\lambda_d = 1 + \sqrt[n]{k}$.

Now suppose Π_n is diagonalizable. Then Π_n admits a basis for $\mathbb{R}^n$ consisting entirely of eigenvectors of Π_n. So any initial vector $\vec{R}_0$ in $\mathbb{R}^n$ can be written as a linear combination of eigenvectors $\vec{R}_0 = \sum\limits_{i=1}^{n} c_i \vec{\lambda}_i$.

Applying Π_n to our initial vector $\vec{R}_0$,

$$\begin{aligned}\vec{R}_1 = \Pi_n \vec{R}_0 = \Pi_n \sum_{i=1}^{n} c_i \vec{\lambda}_i &\\ &= c_1 \Pi_n \vec{\lambda}_1 + \cdots + c_n \Pi_n \vec{\lambda}_n \\ &= c_1 \lambda_1 \vec{\lambda}_1 + \cdots + c_n \lambda_n \vec{\lambda}_n.\end{aligned}$$

Applying Π_n again,

$$\begin{aligned}\vec{R}_2 = \Pi_n^2 \vec{R}_0 = \Pi_n(\Pi_n \vec{R}_0) &= \Pi_n(c_1 \lambda_1 \vec{\lambda}_1 + \cdots + c_n \lambda_n \vec{\lambda}_n) \\ &= c_1 \lambda_1 \Pi_n \vec{\lambda}_1 + \cdots + c_n \lambda_n \Pi_n \vec{\lambda}_n \\ &= c_1 \lambda_1^2 \vec{\lambda}_1 + \cdots + c_n \lambda_n^2 \vec{\lambda}_n.\end{aligned}$$

We can see the pattern now,

$$\vec{R}_t = \Pi_n^t \vec{R}_0 = c_1 {\lambda_1}^t \vec{\lambda}_1 + \cdots + c_n {\lambda_n}^t \vec{\lambda}_n = \sum_{i=1}^{n} c_i \lambda_i^t \vec{\lambda}_i. \tag{10}$$

Verify that indeed,

$$\vec{R}_t = \Pi_n^t \vec{R}_0 = \begin{pmatrix} c_1 {\lambda_1}^t \vec{\lambda}_{1,1} + \cdots + c_n {\lambda_n}^t \vec{\lambda}_{n,1} \\ c_1 {\lambda_1}^t \vec{\lambda}_{1,2} + \cdots + c_n {\lambda_n}^t \vec{\lambda}_{n,2} \\ \vdots \\ c_1 {\lambda_1}^t \vec{\lambda}_{1,n} + \cdots + c_n {\lambda_n}^t \vec{\lambda}_{n,n} \end{pmatrix},$$

where $\vec{\lambda}_{a,b}$ denotes the bth entry of the eigenvector corresponding to λ_a.

Note that with increasing t, only one of the $c_i \lambda_i^t \vec{\lambda}_{i,l}$ terms becomes the dominant term. The dominant term is the one involving the dominant eigenvalue, λ_d and the entries of its corresponding eigenvector, $\vec{\lambda}_{d,i}$. Since contributions of the other terms become negligible in the limiting quotient, we can make the following statement:

$$\vec{R}_t = \Pi_n^t \vec{R}_0 \approx c \lambda_d^t \vec{\lambda}_d = c(1 + \sqrt[n]{k})^t \begin{pmatrix} \left(\sqrt[n]{k}\right)^{n-1} \\ \left(\sqrt[n]{k}\right)^{n-2} \\ \vdots \\ 1 \end{pmatrix} \text{ for } t \gg 1.$$

It follows that

$$\begin{aligned} \frac{\vec{R}_{t,i}}{\vec{R}_{t,i+1}} &= \frac{c(1 + \sqrt[n]{k})^t \vec{\lambda}_{d,i} + \cdots + c_n {\lambda_n}^t \lambda_{n,i}}{c(1 + \sqrt[n]{k})^t \vec{\lambda}_{d,i+1} + \cdots + c_n {\lambda_n}^t \lambda_{n,i+1}} \\ &\approx \frac{c(1 + \sqrt[n]{k})^t \vec{\lambda}_{d,i}}{c(1 + \sqrt[n]{k})^t \vec{\lambda}_{d,i+1}} = \frac{\vec{\lambda}_{d,i}}{\vec{\lambda}_{d,i+1}} = \sqrt[n]{k}, \end{aligned} \tag{11}$$

for $(1 \leq i \leq n-1)$ and $t \gg 1$.

When $n = 2$ (finding square roots), let $c = 0$ and notice that (11) is approximately $-\sqrt{k}$. To satisfy our curiosity from Section 1, if we are to have a sequence generated by (2) converging to $-\sqrt{k}$ then the appropriate initial values are ones which are components of some multiple of the second eigenvector, not a linear combination of the dominant eigenvector and the other. This way, ratios of successive entries of an evolving vector equal the slope of the second eigenvector, $-\sqrt{k}$. If we restrict our choices of initial values to $\mathbb{Q}^2$, then we do not run into this problem of multiple limits.

5 Derivation of the matrix Π_n

In the preceding section we showed that if (8) can be represented as (9), then $\lim_{t\to+\infty} x_{t,i}/x_{t,i+1} = \sqrt[n]{k}$, as desired. It only remains to find the exact form for Π_n. In the analysis above, we showed that long-term time evolution of an initial vector depends heavily on "hitting" $\vec{\lambda}_d$ with its corresponding eigenvalue λ_d, which in turn depends on left-multiplying $\vec{R}_0$ by Π_n. Carrying this calculation out gives

$$\lambda_d\vec{\lambda}_d = \left(1+\sqrt[n]{k}\right)\begin{pmatrix} \left(\sqrt[n]{k}\right)^{n-1} \\ \left(\sqrt[n]{k}\right)^{n-2} \\ \left(\sqrt[n]{k}\right)^{n-3} \\ \left(\sqrt[n]{k}\right)^{n-4} \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \left(\sqrt[n]{k}\right)^{n-1} \\ \left(\sqrt[n]{k}\right)^{n-2} \\ \left(\sqrt[n]{k}\right)^{n-3} \\ \left(\sqrt[n]{k}\right)^{n-4} \\ \vdots \\ 1 \end{pmatrix} + \begin{pmatrix} k \\ \left(\sqrt[n]{k}\right)^{n-1} \\ \left(\sqrt[n]{k}\right)^{n-2} \\ \left(\sqrt[n]{k}\right)^{n-3} \\ \vdots \\ \sqrt[n]{k} \end{pmatrix}.$$

Evidently, Π_n is such that when an initial vector $\vec{R}_0$ is left multiplied by it, returned is the dominant eigenvector plus another vector, $\vec{v}$, along with other negligible terms in the limit that $t \gg 1$. This means that Π_n must be the sum of two matrices acting on the linear combination of eigenvectors that comprises $\vec{R}_0$.

$$\Pi_n\vec{R}_0 = (I_n + \pi_n)\left(\lambda_d\vec{\lambda}_d + \cdots + \lambda_n\vec{\lambda}_n\right).$$

where I_n is the n-dimensional identity matrix. But what is π_n? Consider what action is taken by π_n to return $\vec{v}$ from $\vec{\lambda}_d$. Apparently, π_n is an $n\times n$ matrix such that when it left multiplies a column vector, it has the effect of permuting entries by one place in a cyclic manner while scaling by a factor of $\sqrt[n]{k}$. By inspection, we see that

$$\pi_n = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & k \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

The exact form of Π_n is given by

$$\begin{aligned} \Pi_n &= I_n + \pi_n \\ &= \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & k \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & k \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{pmatrix}. \end{aligned}$$

6 Algebraic properties of the matrix Π_n

We must still confirm diagonalizability of Π_n since most of our case depends upon this property of Π_n. The characteristic polynomial, $P(\lambda)$, of Π_n can be found by computing $\det(\Pi_n - \lambda I_n)$ by expanding in minors along the top row, giving $P(\lambda) = (1-\lambda)^n + (-1)^{n+1}k$. The eigenvalues are had by solving $P(\lambda) = 0$ giving $\lambda_j = 1 - \sqrt[n]{k}\ e^{iJ\pi/n}$ where $1 \leq j \leq n$ and $J = 2j$ when n is even and $2j+1$ when n is odd. The largest of these eigenvalues is $\lambda_d = \lambda_{\frac{n}{2}} = \lambda_{\frac{n-1}{2}} = 1 + \sqrt[n]{k}$ as desired. These n distinct eigenvalues give

n distinct eigenvectors given by

$$\vec{\lambda}_j = \begin{pmatrix} \frac{k}{\lambda_j - 1} \\ \frac{k}{(\lambda_j-1)^2} \\ \frac{k}{(\lambda_j-1)^3} \\ \frac{k}{(\lambda_j-1)^4} \\ \vdots \\ 1 \end{pmatrix},$$

the largest of which corresponds to the eigenvalue λ_d; this vector is

$$\vec{\lambda}_d = \begin{pmatrix} \left(\sqrt[n]{k}\right)^{n-1} \\ \left(\sqrt[n]{k}\right)^{n-2} \\ \left(\sqrt[n]{k}\right)^{n-3} \\ \left(\sqrt[n]{k}\right)^{n-4} \\ \vdots \\ 1 \end{pmatrix}.$$

Since Π gives n distinct eigenvalues and eigenvectors, Π_n is diagonalizable. We are now in the position to make the generalization of (7) as follows

$$\Pi_n^t \vec{R}_0 = \vec{R}_t \text{ gives rise to } \lim_{t\to+\infty} \frac{\vec{R}_{t,i}}{\vec{R}_{t,i+1}} = \sqrt[n]{k} \tag{12}$$

for $1 \leq i \leq n-1$ and $t \in \mathbb{Z}$. Equivalently,

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & k \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{pmatrix}^t \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \\ \vdots \\ x'_n \end{pmatrix}$$

so that $\lim_{t\to+\infty} \frac{x'_i}{x'_{i+1}} = \sqrt[n]{k}$.

This result effectively fulfils the goal of the paper which was to derive a simple recursion that generates a sequence of fractions approximating $\sqrt[n]{k}$ with increasing accuracy.

7 Computation

The result of the previous section satisfies the technical goal of this paper, but it is left to the reader to judge the practicality of this result. Accuracy of an approximation depends on taking powers of an $n \times n$ matrix. This tedious task can be tiresome for even relatively small powers of n and t. So where do we look to find aid in this computation? One could certainly start with the Cayley-Hamilton Theorem, which states that every $n \times n$ matrix over a commutative ring satisfies its own characteristic equation, $P(\lambda) = \det(\lambda I_n - A)$. Applying this theorem to Π_n gives

$$P(\Pi_n) = (I_n - \Pi_n)^n + (-1)^{n+1}k\ I_n = \mathbf{0},$$

where $\mathbf{0}$ is the $n \times n$ matrix consisting entirely of zeros. The binomial theorem then gives

$$\sum_{i=0}^{n} \frac{n!(-1)^i}{i!(n-i)!}\Pi_n^i I_n = (-1)^n k\ I_n.$$

Solving then for Π_n^n gives

$$\Pi_n^n = \sum_{i=1}^{n-1} \frac{n!(-1)^{n-1-i}}{i!(n-i)!}\Pi_n^i + \left[(-1)^{n-1} + k\right] I_n. \tag{13}$$

This is an explicit equation expressing Π_n^n in terms of lower powers of Π_n and I_n. It is useful because if one is able to calculate powers of Π_n up to and including Π_n^{n-1}, then one is able to generate arbitrarily large powers of Π_n iteratively which then can be used to generate arbitrarily close approximations to $\sqrt[n]{k}$.

Let's take a look at the $n = 2$ case. Equation (13) gives

$$\Pi_2^2 = 2\Pi_n + (k-1)I_2, \tag{14}$$

which is an explicit expression of Π_2 in first powers of Π_2 and I_2. Because no higher powers of Π_2 need to be calculated to arrive at (14), arbitrary integer powers of Π_2 are gotten with ease from iterative multiplication and substitution of powers of Π_2. This gives rise to a Fibonacci-like sequence in the exponents of Π_2:

$$
\begin{aligned}
\Pi_2^2 &= 2\Pi_n + (k-1)I_2 \\
\Pi_2^3 &= \Pi_2\Pi_2^2 \\
&= 2\Pi_2^2 + (k-1)\Pi_2 \\
&= (k+3)\Pi_2 + 2(k-1)I_2 \\
\Pi_2^5 &= \Pi_2^3\Pi_2^2 \\
&= (k^2+10k+5)\Pi_2 + 4(k^2-1)I_2 \\
&\vdots \\
\Pi_2^{F_i} &= \Pi_2^{F_{i-1}}\Pi_2^{F_{i-2}}.
\end{aligned}
$$

The reader is encouraged to try this for the $n = 3, 4, 5, ...$ cases to see that once harrowing computations are done to make Π_n^{n-1} known, precise approximate computation soon follows.

References

[1] M. R. Rychlik, Periodic points of the billiard ball map in a convex domain, Journal of Differential Geometry 30 (1989), no. 1, 191-205.

[2] L. Stojanov, Note on the periodic points of the billiard, Journal of Differential Geometry 34 (1991), no. 3, 835-837.

[3] M. P. Wojtkowski, Two applications of Jacobi fields to the billiard ball problem, Journal of Differential Geometry 40 (1994), no. 1, 155-164.

[4] V. Ya. Ivrii, The second term of the spectral asymptotics for a laplace-beltrami operator on manifolds with boundary, Func. Anal. Appl. 14 (2) (1980), 98-106.

[5] Ya. B. Vorobets, On the measure of the set of periodic points of a billiard, Mathematical Notes 55(1994), no. 5, 455-460.

[6] N. Lazrieva, T. Sharia, T. Toronjadze, Semimartingale Stochastic Approximation Procedures and Recursive Estimation, Applied Sciences 22 (1990), no. 2, 91-105.

[7] Luca Heltai, Francesco Costanzo, Variational Implementation of Immersed Finite Element Methods, Central European Journal of Mathematics 32 (1993), no. 4, 72-83.

[8] Steffen Roch, Fractal algebras of discretization sequences, Applied Mathematics Notes 39 (1998), no. 8, 123-140.

[9] Mark Blume, Toric orbifolds associated to Cartan matrices, Excursions in Algebraic and Geometric Topology. 12 (2004), 93-107.

[10] Paul G. Constantine, Eric T. Phipps, A Lanczos Method for Approximating Composite Functions, Conformal Geometry and Dynamics 45 (2008), no. 6, 133-148.